

Manipuler, analyser et visualiser des données grâce aux modules Python de Data Science

RS6763 · RS · Publiée

L'essentiel

Code(s) NSF	110, 313, 326
Formacode(s)	31025, 30812, 31026, 31035
Type d'enregistrement	Enregistrement sur demande
Date d'enregistrement	01/10/2024
Date d'échéance	01/10/2026

Certificateur(s)

API SOCIETY · SIRET : 83381699400010

Partenaires habilités

39 organismes habilités, d'après la liste publiée par France compétences : 39 à former.

Organisme	SIRET	Habilitation
AXIO FORMATION	89250203000026	Former
DATAGONG	91314205500015	Former
DKFORMATIONS	94006274800014	Former
DRS SCHOOL	85398915000018	Former
ECOLE ADR TECH	98382199200010	Former
JEDHA	83872651100044	Former
MAKINA SAPIENS	82286963200031	Former
NATHANIEL BENCHIMOL	52349020900039	Former
NUMGRADE	84336226000013	Former
ORSYS	48276116000019	Former

Liste complète des 39 partenaires habilités : <https://data.aktaconseil.fr/certifications/rs6763-manipuler-analyser-et-visualiser-des-donnees-grace-aux-modules-python-de-data-science/partenaires>

Résumé de la certification

OBJECTIFS ET CONTEXTE

Cette certification s'adresse aux professionnels de la donnée (statisticiens, actuaires, data analysts et data scientists) et aux scientifiques voulant attester leurs compétences en manipulation, analyse et visualisation des données avec les modules Python de Data Science.

COMPÉTENCES ATTESTÉES

Installer dans un environnement virtuel des modules et paquets scientifiques pour manipuler des données

Utiliser un environnement de développement scientifique afin de bénéficier d'une solution intégrée et ainsi être plus rapide et performant

Utiliser des outils collaboratifs afin de travailler à plusieurs data scientists sur les mêmes données

Trouver de nouvelles librairies et juger de leur pertinence en vue de répondre à un besoin particulier

Charger différents formats de données dans un tableau de données indexées ou non en vue de leur appliquer des traitements

Appliquer des fonctions et méthodes sur un tableau de données afin de le nettoyer et de le préparer à un futur traitement

Concevoir des masques pour extraire des données

Filtrer des données selon des plages temporelles afin d'extraire les données correspondantes à un moment précis

Filtrer des données à l'aide d'expressions rationnelles (regex) afin d'extraire les données correspondantes à un motif précis

Changer la fréquence d'échantillonnage de données afin de les ramener sur une échelle adaptée au besoin

Maîtriser les tableaux à n-dimensions (ndarray) et les types de données (entier, réel, chaîne de caractères, date et heure) afin de minimiser l'utilisation de la mémoire

Maîtriser les tableaux à données indexées à une dimension (Serie) et à deux dimensions (DataFrame) afin d'accéder aux fonctionnalités offertes par le module scientifique

Maîtriser l'indexation implicite (indigage par position numérique) afin d'extraire des données et/ou de leur appliquer un traitement

Maîtriser l'indexation explicite (indigage par clé alphanumérique) afin de manipuler plus facilement des données temporelles ou utilisant un identifiant alphanumérique

Faire pivoter des données pour obtenir un tableau croisé dynamique

Maîtriser les index multiples (multi-index) afin d'accéder plus facilement aux données et/ou de simuler des dimensions supplémentaires

Empiler et fusionner des tableaux afin de regrouper des données

Transposer des données (inversion ligne / colonne) pour en faciliter l'analyse, ou les préparer en vue de les empiler ou les fusionner

Trier des données en fonction des indices ou des valeurs du tableau en vue de les exploiter par ordre alphabétique, par nombre croissant ou décroissant, par date

Connaître et utiliser les opérations vectorisées pour gagner en performance

Appliquer des fonctions et méthodes sur une colonne d'un tableau de données afin d'appliquer des traitements mathématiques (nombre, minimum, maximum, moyenne, médiane, centile, écart-type)

Utiliser les fonctions universelles selon une dimension

Regrouper des données en fonction de la valeur de certaines colonnes pour appliquer un calcul

Calculer des moyennes glissantes (lissage des données) en vue de les visualiser

Utiliser les règles de broadcasting pour effectuer des calculs

Appliquer une fonction ou une fonction anonyme aux données pour réaliser des analyses spécifiques

Choisir le bon module ou paquet de visualisation des données afin de répondre au besoin exprimé

Afficher un graphique en 2 dimensions de type $y = f(x)$ à partir d'un jeu de données simple

Identifier le bon type de graphique (histogramme, graphique en colonne, diagramme circulaire, boîte à moustaches, carte thermique ...) en fonction des données à visualiser afin de générer un graphique pertinent

Connaître et utiliser les fonctions et méthodes des tableaux de données afin de générer très facilement des graphiques

Générer des graphiques interactifs afin de permettre à l'utilisateur d'interagir avec le graphique (isolation de données, agrandissement d'une zone, mise en évidence)

Concevoir des graphiques contenant plusieurs tracés afin de regrouper des données différentes, de montrer des corrélations ou des différences, ou de synthétiser une représentation des données

Intégrer les recommandations d'amélioration de l'accessibilité (via l'ajout de titres, étiquettes, labels et légendes, l'amélioration de la lisibilité des textes, l'épaississement des lignes et des bordures, le choix de couleurs contrastées) afin de générer des graphiques accessibles à une personne en situation de handicap

Voies d'accès

PRÉREQUIS À L'ENTRÉE EN FORMATION

Connaître les fondamentaux de la programmation en langage Python

PRÉREQUIS À LA VALIDATION

Avoir acquis les fondamentaux des modules Python de Data Science suite à une formation continue ou à un apprentissage autodidacte

COMPOSITION DU JURY — FORMATION CONTINUE, CANDIDATURE INDIVIDUELLE, VAE

Le jury est composé d'un surveillant (extérieur à l'organisme certificateur) vérifiant l'identité du candidat et s'assurant de l'absence de fraudes durant les épreuves théorique et pratique, d'un professionnel de la Data Science en langage Python (extérieur à l'organisme certificateur) corrigeant les exercices réalisés par le candidat durant l'épreuve pratique et relisant les réponses du candidat à l'épreuve théorique, et d'un président (salarié de l'organisme certificateur ou extérieur) harmonisant les résultats des deux épreuves selon les comptes-rendus du surveillant et du professionnel.

Base légale

Date de décision	01/10/2024
Durée de l'enregistrement	2 ans
Date d'échéance de l'enregistrement	01/10/2026
Date de dernière délivrance possible	01/04/2027

Pour plus d'informations

Fiche officielle (France Compétences)

<https://www.francecompetences.fr/recherche/rs/6763/>

Entrées et sorties de formation

Entrées en formation	75
Parcours menés à terme	58
Taux d'achèvement	77 %
Période	09/2025 – 09/2026

Entrées et sorties — Caisse des Dépôts (open data), tous organismes confondus.

Demande CPF — formations engagées

Dossiers CPF engagés	83
Montant total engagé	137 968 €
Coût moyen par dossier	1 662 €
Durée moyenne	90 h
Période	07/2025 – 08/2026
France compétences	86 %
Reste à charge titulaire	10 %
Entreprise	2 %
OPCO	1 %
France Travail	1 %
Compte d'engagement citoyen	0 %

Formations engagées via le CPF — Caisse des Dépôts (open data), tous organismes confondus.